Eur. J. Biochem. 267, 5085–5092 (2000) © FEBS 2000

Highly potent inhibitors of human cathepsin L identified by screening combinatorial pentapeptide amide collections

Achim Brinker[1,2,*], Ekkehard Weber[3], Dieter Stoll[2,4], Jürgen Voigt[1], Annett Müller[5,†], Norbert Sewald[5], Günther Jung[4], Karl-Heinz Wiesmüller[6] and Peter Bohley[1]

[1]*Physiologisch-chemisches Institut and* [2]*Naturwissenschaftliches und Medizinisches an der Universität Tübingen, Germany;* [3]*Institut für Physiologische Chemie der Universität Halle-Wittenberg, Germany;* [4]*Institut für Organische Chemie, Institut der Universität Tübingen, Germany;* [5]*Universität Bielefeld, Fakultät für Chemie, Germany;* [6]*EMC microcollections GmbH, Tübingen, Germany*

By screening a combinatorial pentapeptide amide collection in an inhibition assay, we systematically evaluated the potential of 19 proteinogenic amino acids and seven nonproteinogenic amino acids to serve as building blocks for inhibitors of human cathepsin L. Particularly efficient were aromatic, bulky, hydrophobic amino-acid residues, especially leucine, and positively charged residues, especially arginine. Building blocks for potential inhibitory peptides were combined by random selection from their activity pattern. This random approach for the design of inhibitors was introduced to compensate for the inaccuracy induced by shifted docking of combinatorial compound collections at the active center of cathepsin L. Thereby, we obtained structurally defined pentapeptide amides which inhibited human cathepsin L at nanomolar concentrations. Among the most potent novel inhibitors, one peptide, RKLLW-NH_2, shares the amphiphilic character of the nonamer fragment VMNGLQNRK of the autoinhibitory, substrate-like, but reverse-binding prosegment of human cathepsin L which blocks the active center of the enzyme. Obviously, RKLLW-NH_2 carries the functions that are important for enzyme–peptide interaction in a condensed form. This hypothesis was confirmed by structure–activity studies using truncated and modified pentapeptides.

Keywords: cathepsin L inhibitors; synthetic peptide library.

Cathepsin L (EC 3.4.22.16) is one of the major papain-like cysteine proteases expressed in animal tissues [1,2]. It is ubiquitously distributed in the lysosomes of all cells where it has a prominent role in nonspecific intracellular protein breakdown [2]. Cathepsin L is synthesized as an inactive proenzyme preventing premature proteolytic activity. An N-terminal 96-amino-acid protein extension, the prosegment, serves as a potent and specific macromolecular inhibitor of immature procathepsin L [3,4]. High levels of the proenzyme are secreted by nontransformed cells, such as osteoclasts, fibroblasts and macrophages, as well as by malignantly transformed cells [5]. Conversion to the mature enzyme occurs intracellularly in lysosomes at pH 3.0–3.5 by autocatalytic removal of the prosegment [5,6], whereas extracellularly at pH 5.5–6.0 maturation is supported by negatively charged matrix surfaces [5–7] or metalloproteases [5]. Because of its high activity against a wide variety of protein substrates [5], mature cathepsin L has physiologically important extracellular functions [8–11]. Human cathepsin L is of particular interest because of its suggested involvement in the pathogenesis of rheumatoid arthritis [12], as well as in tumor invasion and metastasis [13]. Inhibition of the enzyme is expected to be a promising anticancer strategy [13–16]. Therefore, human cathepsin L is an interesting target for drug design and potent low molecular mass inhibitors are expected to be good candidates for therapeutic agents.

Crystal structures for human procathepsin L [17] and the mature form of cathepsin L [18] have been published. The enzyme shares the common fold of the papain superfamily. Its two domains are arranged 'V'-like and enclose the central substrate-binding cleft. The hydrophobic S_2-binding site forms the only distinct pocket [19]. Accordingly, the substrate specificity of cathepsin L seems to be largely determined by bulky hydrophobic amino acids in the P2 position of substrate proteins [20,21]. The inhibitory mechanism of the prosegment is common to all members of the papain family [4]. By interacting with the substrate binding cleft in the reverse orientation compared with natural protein substrates, the otherwise substrate-like prosegment blocks the fully established reactive center of the protease without being cleaved.

Diverse structure-based strategies using large fragments and partial sequences of macromolecular inhibitors of cysteine proteases [3,22–24], as well as strategies involving the testing of a limited range of individual substrate homologs [5] have been followed to identify potent inhibitors of members of the cathepsin family of proteases. To date, however, a comprehensive, nonstructure based, and therefore unbiased approach has yet to be reported. Combinatorial chemistry offers nonstructure-based strategies for the study of molecular interactions [25–28]. Thus, the use of combinatorial compound collections is a promising approach to understand the structural

Correspondence to K.-H. Wiesmüller, EMC microcollections GmbH, Sindelfinger Str. 3, 72070 Tübingen, Germany.
E-mail: wiesmueller@microcollections.de

Abbreviations: O, defined position in pentapeptide amide collections; O/X_4-NH_2, collection characterized by one defined amino acid in its position; X, randomized position in peptide collections characterized by a mixture of amino-acid residues A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y.

Enzyme: cathepsin L (EC 3.4.22.16).

Present address: MPI für Biochemie, 82512 Martinsried, Germany.
†*Present address*: EVOTEC BioSystems AG, Hamburg, Germany.

(Received 9 March 2000, accepted 13 June 2000)

requirements for potent inhibitors of human cathepsin L. In the search for novel inhibitory peptides of human cathepsin L, we first determined the sequence-dependent features of proteinogenic and seven nonproteinogenic amino acids by testing a series of 130 combinatorial peptide amide collections. With one defined position in each collection each individual $O/X_4\text{-}NH_2$ mixture contained a calculated number of 19^4 different peptides. In order to obtain peptide inhibitors, we combined amino acids at all sequence positions that were found to be conducive to inhibition in the respective sequence positions.

EXPERIMENTAL PROCEDURES

Synthesis and analysis of defined peptides and pentapeptide amide collections

Synthetic pentapeptide amide collections, as well as individual pentapeptide amides, were prepared by fully automated solid-phase peptide synthesis using Fmoc/tBu chemistry and Rink amide 4-methylbenzhydrylamine polystyrene resins as described recently [29]. In brief, in the case of peptide amide collections the introduction of randomized sequence positions (X) was performed in a double coupling step with premixed equimolar mixtures of Fmoc-L-amino acids that were used in an equimolar ratio with respect to the coupling sites of the resins. For coupling of defined sequence positions (O), a fivefold molar excess of single Fmoc-L-amino acids was used. An optimized diisopropylcarbodiimide/1-hydroxybenzotriazole method was used to synthesize the collections to yield close to equimolar mixtures [30]. Defined peptides were prepared by multiple peptide syntheses [31]. *N*-Acetylation was performed by reacting the N-terminally deprotected resin-bound peptide with a 1 : 1 mixture of diisopropylethylamine [1.5 M in dichloromethane/dimethylformamide (1 : 1)] and acetic anhydride [1.5 M in dichloromethane/dimethylformamide (1 : 1)] for 30 min. To synthesize pentapeptides (C-terminal carboxy groups), 2-chlorotritylchloride/1%-divinylbenzene–polystyrene resins (Novabiochem, Bad Soden, Germany) were used. The identity of defined peptides was confirmed by electrospray ionization mass spectrometry (ESI-MS) and the purity was > 80% as determined by HPLC. The amino-acid composition in the defined position and the random sequence positions of $O/X_4\text{-}NH_2$ peptide mixtures was determined by pool sequencing [32] and ESI-MS [33]. Deviations from equimolar representation of the amino acids in randomized sequence positions were found to be within the error limits of the analytical methods. Fmoc-protected amino acids β-leucine, homo-β-leucine and homo-β-isoleucine were prepared by a modified Arndt-Eistert synthesis [34].

Human cathepsin L

Human procathepsin L was purified from the culture medium of the human non-small-cell lung cancer cell line EPLC 32M1 as described elsewhere [35] and converted to mature and fully active single chain cathepsin L on a negatively charged surface. Procathepsin L was applied to a carboxymethyl–Sephadex C 50 column equilibrated with 0.01 M sodium acetate (pH 5.5). Conversion to the mature enzyme occurred during elution with a linear gradient of NaCl (0–0.5 M). Cathépsin L-containing fractions were concentrated and 0.5 μL 0.05 M $HgCl_2$ and 0.5 μL 0.1 M EDTA were added to 100 μL aliquots of the enzyme stock solution (118 $\mu g \cdot mL^{-1}$). These aliquots were stored at −70 °C until use.

Cathepsin L inhibition assay

The stock solution of human cathepsin L was diluted 10 000-fold with reactivation buffer (0.1 M NaAc, pH 5.5; 0.01% Triton-X 100; 5 mM EDTA; 5 mM dithiothreitol) and incubated at 0 °C for 30 min. The inhibition assay was performed in 96-well assay plates. Aliquots of the activated enzyme solution (50 μL) were pre-incubated in the absence or presence of putative peptide inhibitors (50 μM) for 30 min at 0 °C. The reaction was started by the addition of substrate (10 μL of Z-Phe-Arg-7-amino-4-methylcoumarine, 10 μM in reactivation buffer) [36]. After 20 min at 37 °C, the reaction was terminated by the addition of 100 μL stop solution (1 M NaAc, 200 mM sodium-monochloroacetate, pH 3.0) and fluorescence intensities were measured using a luminescence spectrometer (type LS50B, PerkinElmer, Überlingen). All tests were performed in triplicate. The pentapeptide mixtures were solubilized in aqueous dimethylsulfoxide (1% v/v), defined peptides in aqueous dimethylsulfoxide (5% v/v). Extremely hydrophobic individual peptides were solubilized in aqueous dimethylsulfoxide (5% v/v) by the addition of Triton-X 100 (0.01% v/v). Enzyme activities were measured in the presence of the $X_5\text{-}NH_2$ collection or of $O/X_4\text{-}NH_2$ collections and were expressed as relative enzyme activities in percent values of the maximal activity achieved in the absence of peptides (100%).

RESULTS

Screening of a combinatorial peptide amide library

To identify promising amino-acid building blocks of human cathepsin L inhibitors, a combinatorial pentapeptide amide collection was investigated in an inhibition assay against the enzyme. Collections represented by the formulas $OXXXX\text{-}NH_2$, $XOXXX\text{-}NH_2$, $XXOXX\text{-}NH_2$, $XXXOX\text{-}NH_2$ and $XXXXO\text{-}NH_2$ were prepared. They differed in the position of the defined amino acid represented by O. Randomized positions X reflected a mixture of 19 (cysteine excluded) of the 20 genetically encoded amino acids (X = A,D,E,F,G,H,I,K,L,M, N,P,Q,R,S,T,V,W,Y). Defined positions (O) were represented by one of these 19 proteinogenic amino-acid residues or by one of the seven nonproteinogenic amino acids citrullin, hydroxyproline, tetrahydroisochinoline-3-carbonic acid, α-pyridylalanine, α-naphthylalanine, *p*-nitrophenylalanine or homophenylalanine (B). Thus, each of the five sets of $O/X_4\text{-}NH_2$ collections contained 26 different subsets, each representing 19^4 individual peptides, for a grand total of $26 \times 19^4 = 3.4 \times 10^6$ individual peptides.

In addition, the completely randomized collection $X_5\text{-}NH_2$ was prepared and tested as a standard. $X_5\text{-}NH_2$ represented the average inhibitory effect of all peptides present in the combinatorial pentapeptide amide collection made from 19 proteinogenic amino acids. Individual $O/X_4\text{-}NH_2$ partial collections represented the inhibitory effect of the subset of pentapeptide amides that contained a distinct defined amino acid at a particular sequence position. Amino acids O useful as building blocks of inhibitors are expected to increase the inhibitory effect of $O/X_4\text{-}NH_2$ collections compared with $X_5\text{-}NH_2$ and to cause a stronger reduction of enzyme activity.

In the inhibition assay, all peptide amide collections were tested close to the IC_{50} concentration of $X_5\text{-}NH_2$. At a final concentration of 136 μM, $X_5\text{-}NH_2$ reduced the enzyme activity to 48%. Therefore, 48% enzyme activity was defined as the reference value and was subtracted from all enzyme activities

achieved in the presence of O/X_4-NH_2 collections. The results from testing 130 O/X_4-NH_2 collections are given in Fig. 1. Amino acids promoting inhibition showed negative relative enzyme activities. Defined amino acids (O) which were regarded as favorable for inhibition showed relative enzyme activities of $< -10\%$.

Significant differences in the inhibitory activities of O/X_4-NH_2 collections were found to be dependent both on the defined amino-acid residue and their positions within the sequence. Therefore, it was possible to estimate the potential of 19 genetically encoded and seven nongenetically encoded amino acids to serve as building blocks for inhibitors of human cathepsin L. The bulky aliphatic amino acids leucine and methionine, the aromatic amino acids phenylalanine, tyrosine and tryptophan, as well as the positively charged amino acids arginine and lysine considerably improved the inhibitory effects when representing O within O/X_4-NH_2 collections. In contrast, the negatively charged amino acids aspartate and glutamate, the small aliphatic amino acids alanine and glycine, and the amino acids proline, asparagine and glutamine were nonfavorable for inhibition when representing O in O/X_4-NH_2 collections. Interestingly, leucine and methionine promoted, whereas isoleucine hindered, the inhibitory activity of O/X_4-NH_2 collections (Fig. 1). The introduction of the phenylalanine analogs *p*-nitrophenylalanine or homophenylalanine in O/X_4-NH_2 collections had no influence. The introduction of other nonproteinogenic amino acids such as citrullin, hydroxyproline, tetrahydroisochinoline-3-carbonic acid, α-pyridylalanine and α-naphtylalanine caused a considerable decrease in inhibitory activities of O/X_4-NH_2 collections.

The effects described so far were not strongly influenced by the positions of the defined amino acids in the pentapeptide amides. However, in position 3, the favorable effect of leucine was more pronounced than in all other positions. Phenylalanine was the second most favorable amino acid in this position, but was considerably less effective than leucine. Basic amino acids showed no pronounced effect at position 3. In position 5, tryptophan instead of leucine was particularly conducive to inhibition.

Deduction of individual pentapeptide amide inhibitors

The cut-off value for considering amino acids as building blocks of putative cathepsin L inhibitors was set at -10% relative enzyme activity in the presence of the respective O/X_4-NH_2 collections. The mean activity values of nearly all O/X_4-NH_2 collections that reduced the relative enzyme activity to at least -10% (Fig. 1) were significantly different from the

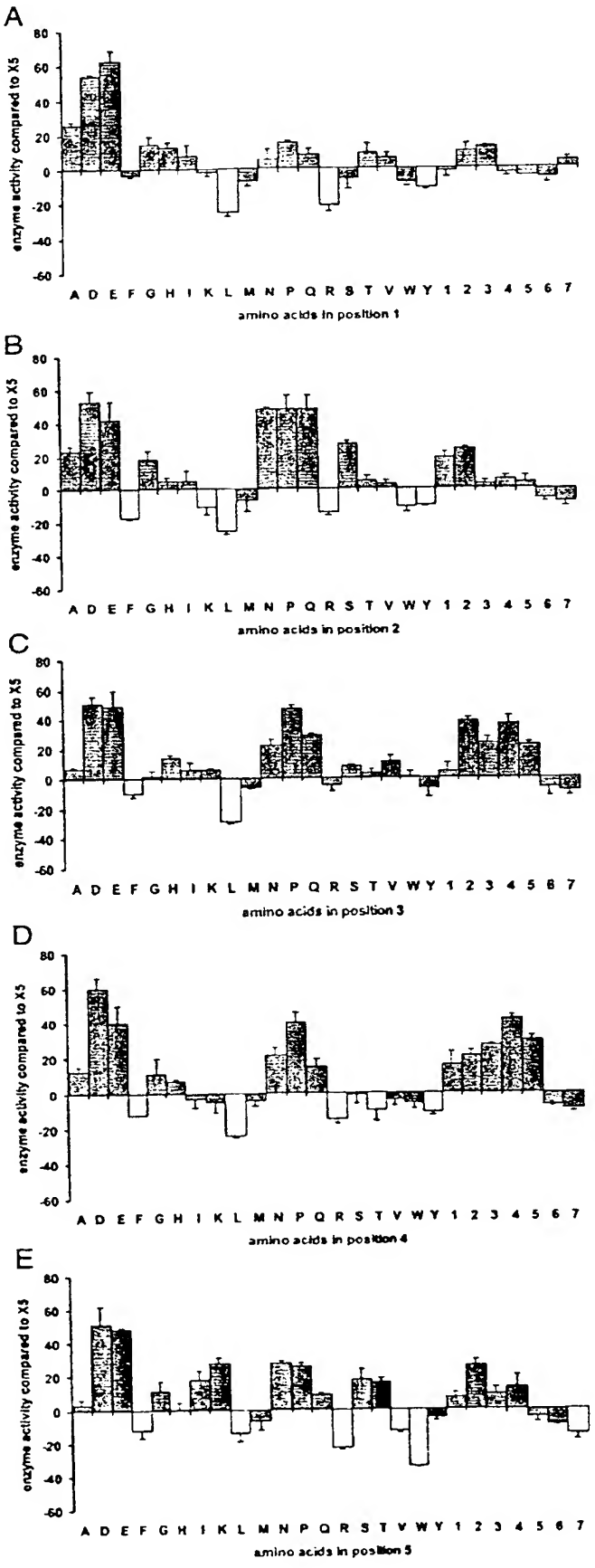

Fig. 1. One-hundred and thirty O/X_4-NH_2 collections and the X_5-NH_2 collection were tested for inhibition of human cathepsin L at concentrations of 136 μM. The defined positions (O) contained 19 proteinogenic amino acids and seven unusual amino acids. (1) citrullin, (2) hydroxyproline, (3) tetrahydroisochinoline-3-carbonic acid, (4) α-pyridylalanine, (5) α-naphthylalanine, (6) *p*-nitrophenylalanine, and (7) homophenylalanine. X_5-NH_2 reduced the enzyme activity to 48%. The effects of O/X_4-NH_2 collections are shown as enzyme activitites relative to the baseline value of 48%. Defined amino acids (O) promoting inhibition in O/X_4-NH_2 collections relative to the average effect achieved by X_5-NH_2 show negative relative enzyme activities. Amino acids that were regarded favorable effected relative enzyme activities of $< -10\%$ and are marked by empty bars. (A–E) Results for all five groups of pentapeptide amide collections OXXXX-NH_2, XOXXX-NH_2, XXOXX-NH_2, XXXOX-NH_2 and XXXXO-NH_2 with defined amino acids at the five sequence positions.

mean activity values of the X_5-NH_2 collection at the 5% significance level (Student's *t*-test) with the exception of the O/X_4-NH_2 collection carrying Thr at position 4. Nevertheless, for the design of individual peptides, Thr at position 4 was considered to increase the diversity of putative cathepsin L inhibitors. The amino acids that reduced the relative enzyme activity to at least −10% are listed in Table 1. Amino acids with properties such as bulky hydrophobic, aromatic and basic were found to be active at different positions.

The results obtained with the peptide collections used in this study could be influenced by shifted docking of O/X_4-NH_2 pentapeptide amides. This is indicated by residues repetitively identified as favorable for inhibition at different positions (Table 1) and complicates the directed selection of amino acids from the screening results. To evaluate these assumptions, amino-acid residues for putative pentapeptide inhibitors were randomly selected from the amino acids listed in Table 1. Of numerous possible combinations, an initial set of 76 sequences was defined and synthesized (Table 2). Forty-nine peptides obeyed the rules for amino-acid residues and positions given in Table 1. Eighteen peptides contained tyrosine, arginine or lysine instead of phenylalanine or leucine at position 3. However, they were not among the best inhibitors, and neither were two control peptides with glutamine in position 5 (Table 2). Three peptides with tryptophan in position 4 were prepared (Table 2; 4, 25, 32) and, unexpectedly, peptide number 4, RKLWL, proved to be one of the most active inhibitors (Table 3). Valine replacing leucine at position 5 (peptide RKLWV) was nonfavorable for inhibition. Four negative controls were synthesized containing negatively charged amino acids (Table 2, numbers 65, 71, 72 and 76).

In summary, peptide amides, such as RKLLW-NH_2, that were composed solely of the favorable amino acids listed in Table 1 showed strong activity, whereas negatively charged pentapeptide amides, such as EEEEE-NH_2, did not influence the enzyme activity up to a concentration of 180 μM (Fig. 2). Several inhibitors with IC_{50} values in the submicromolar range could be identified. In accordance with the results obtained by screening of the O/X_4-NH_2 collections, substitution of one residue within the RKLWL-NH_2 sequence by a negatively charged residue caused a considerable loss in the inhibition potential (data not shown), e.g. the IC_{50} value of RKLWD-NH_2 was 18 times that of RKLWL-NH_2.

Modifications of model peptide RKLLW-NH_2

From the most efficient inhibitors identified, the peptide amide RKLLW-NH_2 was chosen as a model peptide for further investigations. It contains the structural properties which supported inhibition of human cathepsin L by pentapeptide amides. The K_i of RKLLW-NH_2 was determined to be 130 nM. Whereas RKLLW-NH_2 reduced the enzyme activity to 34% at a concentration of 0.9 μM (Fig. 2), the inverted peptide WLLKR-NH_2 showed no inhibitory activity (data not shown). To identify a functional minimal core structure of RKLLW-NH_2, homologous dipeptide, tripeptide, and tetrapeptide amides were tested

Table 1. Amino acids favorable for the design of cathepsin L inhibitors. Amino acids O in O/X_4-NH_2 collections that reduced the relative enzyme activity of at least 10% are listed in their respective sequence positions.

Position	1	2	3	4	5
	L	L	L	L	W
	R	F	F	R	R
	Y	R		F	L
		K		Y	V
		W		T	F
		Y			

Table 2. Defined pentadecapeptide amides for evaluation of the screening results obtained with O/X4-NH2 collections. B, homophenylalanine.

Peptide	Sequence	Activity (%)	Peptide	Sequence	Activity (%)
1	RKLLW	16	39	LWLFW	69
2	LLLRW	17	40	LLYTW	70
3	LLLTR	17	41	YFLLR	70
4	RKLWL	18	42	RLLYW	71
5	RKLFL	21	43	YFLTF	71
6	LKFTR	22	44	LLYTB	75
7	LFLRL	22	45	LRFTF	75
8	LFLTR	24	46	LKFTF	76
9	LKLLW	25	47	RRFYV	76
10	RLLLW	25	48	LKLFW	78
11	LRLLW	27	49	RFLRW	79
12	LWLFL	27	50	LWRFW	80
13	YWLLR	27	51	LWFRQ	80
14	YKLLR	28	52	YWYYL	80
15	LLLTL	33	53	YFYTW	81
16	RWLYL	35	54	LKYFL	82
17	RFLYR	36	55	LYKFF	83
18	LLYTR	38	56	YKYLW	84
19	LLLLR	38	57	RFFRB	85
20	LLLTW	38	58	RRYRR	87
21	LWFFW	39	59	LKLFF	88
22	RWLTL	39	60	RKYRR	91
23	LLLTB	42	61	LKLRV	91
24	YYLLR	44	62	RFFRR	93
25	RKLWF	48	63	YWLTW	95
26	LLFRW	48	64	LKFRF	95
27	LLFLW	48	65	EDEDE	96
28	LWLLW	48	66	YYYYW	96
29	LLYLW	53	67	RFYRW	98
30	YWFTF	57	68	RFYRL	98
31	LKLFF	58	69	RWFRW	98
32	RKLWV	59	70	YWLLR	101
33	YLYLF	60	71	DEDED	102
34	RRLTW	62	72	EEEEE	103
35	LLLLW	63	73	LKFRB	104
36	YLLFW	63	74	LFYRL	104
37	RRYLB	67	75	RKLRQ	105
38	LKFTF	68	76	EQPPK	112

Table 3. IC_{50} value of four newly defined potent pentapeptide amide inhibitors of cathepsin L. Replacement of C-terminal leucine by aspartic acid reduces activity by a factor of 18.

Sequence	IC_{50} (μM)
LLLTR-NH_2	0.5
RKLLW-NH_2	0.6
LFLTR-NH_2	0.8
RKLWL-NH_2	0.8
RKLWD-NH_2	14.0

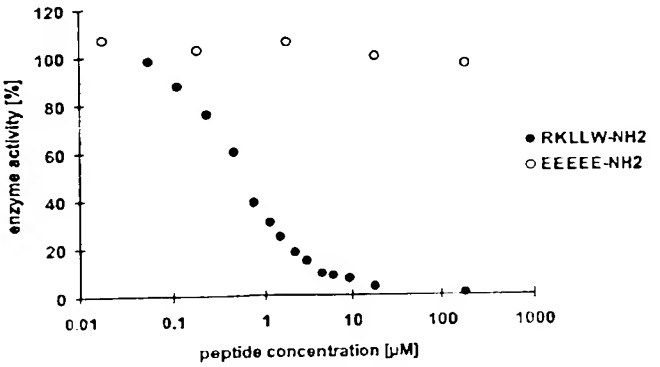

Fig. 2. Titration of the two individual pentapeptide amides RKLLW-NH_2 and EEEEE-NH_2 in a cathepsin L inhibition assay. RKLLW-NH_2 consists of favorable amino acids only and shows an IC_{50} of 0.6 μM, whereas the highly negatively charged peptide EEEEE-NH_2 contains only the nonfavorable glutamate and has no effect on enzyme activity up to a concentration of 180 μM.

for their inhibitory efficiency. However, none of the truncations of RKLLW-NH_2 showed comparable inhibition at a concentration of 0.9 μM (data not shown). Investigations of terminally modified RKLLW-NH_2, such as Ac-RKLLW-NH_2 and RKLLW-COOH, indicated that the positively charged N-terminus promotes the enzyme–peptide interactions, whereas a negatively charged C-terminus is nonfavorable (Fig. 3A).

Successive substitutions of all five amino acids in RKLLW-NH_2 for alanine (alanine scan) were carried out to evaluate the contribution of the distinct amino-acid residues to the activity of the pentapeptide. While substitutions of most residues in RKLLW-NH_2 for alanine had no striking effects on the activity of the resulting pentapeptide amide, the substitution of leucine in position three strongly decreased the activity of this particular inhibitor of cathepsin L (Fig. 3B). Substitutions of

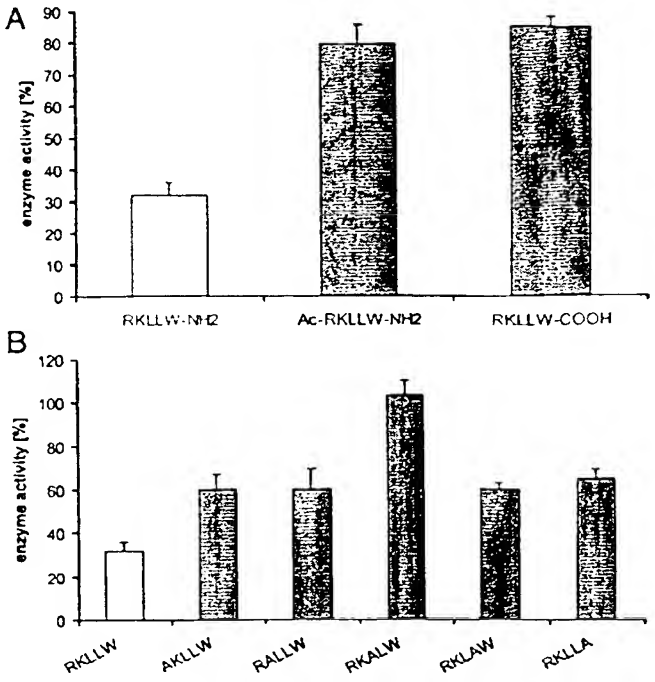

Fig. 3. Influences of N-terminal and C-terminal modifications of RKLLW-NH_2 evaluated by testing Ac-RKLLW-NH_2 and RKLLW-COOH for inhibition of human cathepsin L (A) and Ala-scan of all sequence positions of RKLLW-NH_2 to evaluate the contribution of each individual amino acid to inhibition (B). All pentapeptide amides were tested at a concentration of 0.9 μM.

one L-amino acid for the respective D-amino-acid residue were prepared as were all D-peptides to investigate the stereospecificity of the peptide–enzyme interaction. For all D or individual D substitutions, at a concentration of 0.9 μM (data not shown),

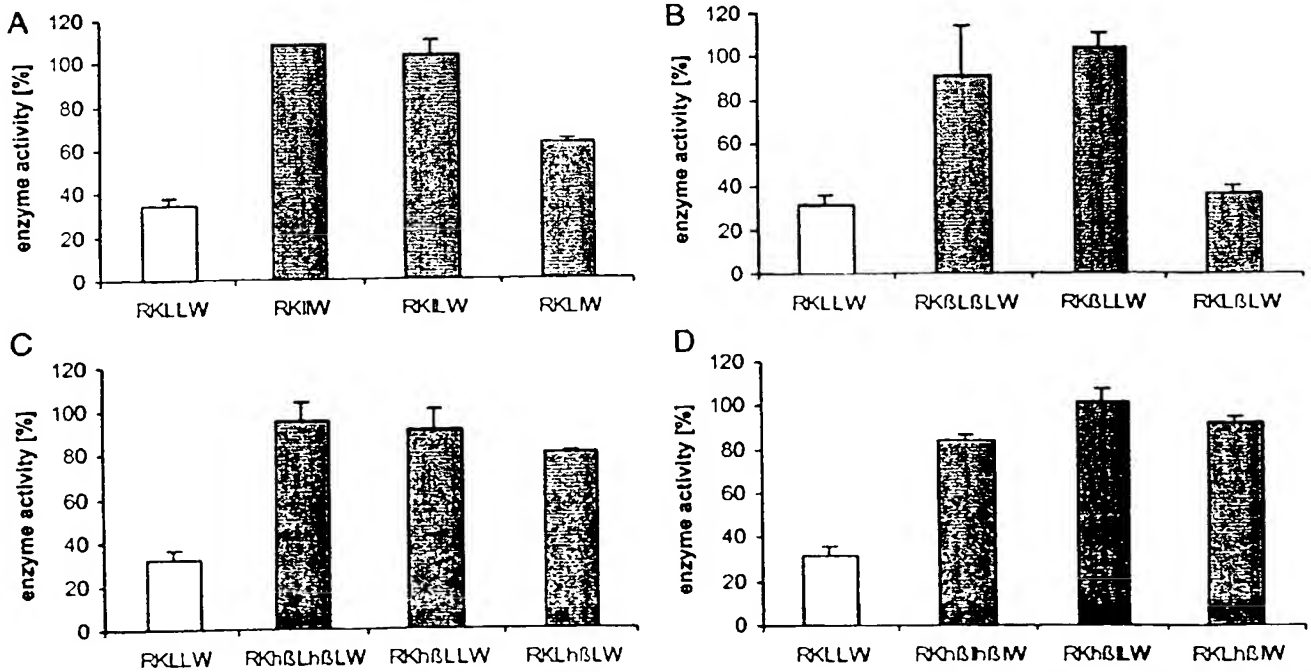

Fig. 4. Substitution of both leucine residues of RKLLW-NH_2 for (A) isoleucine, (B) β-leucine (βL), (C) homo-β-leucine (hβL), and (D) homo-β-isoleucine (hβI). All peptides were tested at a concentration of 0.9 μM.

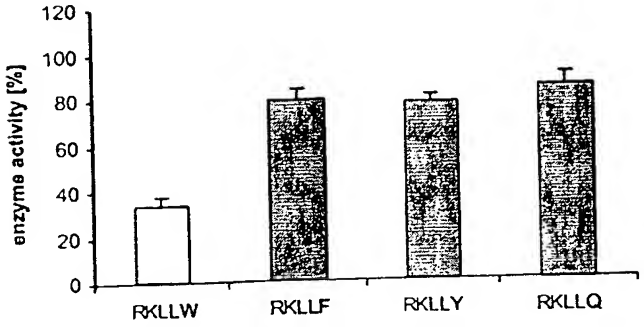

Fig. 5. Investigations of RKLLW-NH_2-derivatives with substitutions in position 5 to further evaluate the contribution of Trp5 to the inhibitory activity of the model peptide. All peptides were tested at a concentration of 0.9 μM.

no inhibitory activity was observed. Reversing the N-terminal amino acids RK to KR had only a marginal influence on the inhibitory effect of the peptide.

The striking impact of Leu3 in RKLLW-NH_2 found in the alanine scan, was further confirmed by simultaneous and successive substitutions of leucine residues for isoleucine (Fig. 4A) and β-leucine (Fig. 4B). In both cases, a pronounced decrease in the biological activity of the peptide was observed. Whereas RKLIW-NH_2 and RKLβLW-NH_2 showed considerable inhibitory effects at a concentration of 0.9 μM, RKIIW-NH_2 and RKILW-NH_2 as well as RKβLβLW-NH_2 and RKβLLW-NH_2 had no influence at all on enzyme activity. Homo-β-leucine was also substituted for each individually, as well as both simultaneously in the case of the leucine residue in RKLLW-NH_2. Neither RKLhβIW-NH_2, RKhβIhβIW-NH_2 nor RKLβhIW-NH_2 showed significant inhibitory effects at 0.9 μM (Fig. 4C). Similarly, the introduction of homo-β-isoleucine completely abolished the inhibitory effects of modified pentapeptide amides at a concentration of 0.9 μM (Fig. 4D). Replacement of Trp5 in RKLLW-NH_2 by phenylalanine or tyrosine or by glutamine considerably reduced the inhibitory activity (Fig. 5).

DISCUSSION

A combinatorial pentapeptide amide collection, O/X_4-NH_2, was used to identify the bulky hydrophobic amino acid leucine, the aromatic amino acids phenylalanine, homophenylalanine, tyrosine and tryptophan, and also the basic amino acids arginine and lysine as potential building blocks for the design of inhibitors against human cathepsin L. The discovery of leucine and the aromatic amino acids is in good agreement with the substrate specificity of cathepsin L. Primary cleavage sites in physiologically relevant protein substrates, such as glucagon and the insulin β chain, are largely determined by bulky hydrophobic and aromatic amino acids in the S2 site [20,21]. Potent reversible peptide inhibitors leupeptin (Ac-Leu-Leu-Arg-CHO), peptide aldehyde semicarbazones and peptidyl-methylketones [5] contain leucine and aromatic amino acids.

Calculations of the electrostatic potential distribution on the enzyme surface revealed a strongly negative potential lobe surrounding the whole substrate binding site of human cathepsin L. The negative potential lobe and the negatively charged enzyme surface in binding distance around the substrate binding cleft, immediately explained the overall favorable effect of positive charges on the inhibitory properties of peptide inhibitors. This concerns K/X_4-NH_2 or R/X_4-NH_2 collections with defined basic residues as well as defined peptides with multiple positive charges in side chains and unprotected N-termini, such as RKLLW-NH_2 or RKLWL-NH_2. Negative charges on peptides in the side chains or at the C-terminus should accordingly have strongly adverse overall effects on inhibition activities.

In combinatorial O/X_4-NH_2 collections, hydrophobic, aromatic and positively charged side chains of defined (O) amino acids appeared to support the inhibition activities independent of their relative sequence position (Table 1). However, leucine and, to a lesser extend, phenylalanine at position 3 where basic amino-acid residues did not show considerable inhibitory activity was extraordinarily effective. The obvious importance of the third sequence position was further corroborated by studies involving modifications of the model inhibitor RKLLW-NH_2. Nevertheless, the general tendency indicated a heterogenous binding of the O/X_4-NH_2 collections relative to the enzyme enabeled by shifted docking of O/X_n-NH_2 collections. Therefore, it was not promising to deduce the optimal relative sequence of the three types of building blocks merely from the most striking results from the inhibition assays. Even though, by arbitrary combination of the favorable bulky hydrophobic, aromatic and basic amino acids, we obtained several individual pentapeptide amides with inhibitory effects against human cathepsin L at nanomolar concentrations. Thus the new compounds (Table 3) rank among the best noncovalently binding low-molecular mass inhibitors of human cathepsin L [5,22]. The most effective inhibitors do not share a common sequence motif. Peptides like RKLLW-NH_2 and RKLWL-NH_2 accumulate three positive charges in their N-terminal regions, whereas peptides like LLLTR-NH_2 or LFLTR-NH_2 contain a basic residue at the C-terminal position and hydrophobic residues in the N-terminal part. These peptides obviously represent different lead structures, which is in good agreement with the broad substrate specificity of cathepsin L [5]. Leucine can be substituted for phenylalanine without substantially altering the inhibitory activity of the peptide. Peptides LFLTR-NH_2 and LLLTR-NH_2 reflect observations made with the reverse oriented inhibitory prosegment. In the mutated form Leu78p could replace wild-type Phe78p and still fit into the S2 pocket [17].

One particularly potent inhibitor, RKLLW-NH_2, compris positively charged and neutral amino acids in a manner similar to the reverse binding nonamer prosegment sequence VMNGLQNRK (residues 74–82) which is located in the substrate binding site of mutated human procathepsin L [17]. RKLLW-NH_2 shares the amphiphilic character of the nonamer fragment as well as a bulky hydrophobic leucine anchor for insertion into the S2 pocket. The three positive charges of RKLLW-NH_2 interact strongly with the negative surface potential around the binding cleft of human cathepsin L and particularly with the rather undefined S3/S4 binding region. Interestingly, the prosegment contains an extraordinary accumulation of five positive charges in residues 81–87 (RKPRKGK) which confirm the importance of basic building blocks for cathepsin L inhibitors in sites other than S1′ [36]. The prosegment-based model is in agreement with our experimental data confirming the prominent leucine anchor at the S2-binding pocket opposed to Leu4 that points toward the solvent.

Interestingly, fragments taken from conserved sequences of cystatins, macromolecular inhibitors of cathepsin L and other papain-like cysteine proteases, have been tested in the course of rational, structure-based approaches for their capacity to act as enzyme inhibitors [21,22]. However, these isolated fragments

adopted a binding mode completely different to the homologous sequences in the macromolecular prosegment [22]. This has been shown for cathepsin B with overlapping pentadecapeptides that inhibit the enzyme in the micromolar range [37] and for irreversible inhibitors of human cathepsin B containing fragments of the cathepsin B prosegment [23]. Synergic to structure-based inhibitor design combinatorial compound collections applied in this study offer a rapid method for selection of building blocks for efficient reversible inibitors of human cathepsin L as shown for the lead structure RKLLW-NH_2. Results from screening O/X_4-NH_2 collections, show that from a total of 130-amino-acid residue-sequence position combinations nine amino acids were positively ranked for one or several sequence positions. The fuzzy positioning was interpreted as shifted docking of amino acids O in O/X_4-NH_2 collections and was taken into consideration by the unvalued selection of the building blocks for novel inhibitors of human cathepsin L.

ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (SFB 323, SFB 510).

REFERENCES

1. Kirschke, H., Langner, J., Wiederanders, B., Ansorge, S. & Bohley, P. (1977) Cathepsin L. A new proteinase from rat-liver lysosomes. *Eur. J. Biochem.* **74**, 293–301.
2. Turk, B., Turk, V. & Turk, D. (1997) Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol. Chem.* **378**, 141–150.
3. Carmona, E., Dufor, E., Plouffe, C., Takebe, S., Mason, P., Mort, J.S. & Mènard, R. (1996) Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases. *Biochemistry* **35**, 8149–8157.
4. Cygler, M. & Mort, J.S. (1997) Proregion structure of members of the papain superfamily. Mode of inhibition of enzymatic activity. *Biochimie* **79**, 645–652.
5. Kirschke, H., Barrett, A.J. & Rawlings, N.D. (1995) Proteinases 1: lysosomal cysteine proteinases. *Protein Profile* **2**, 1581–1643.
6. Ménard, R., Carmona, E., Takebe, S., Dufour, E., Plouffe, C., Mason, P. & Mort, J.S. (1998) Autocatalytic processing of recombinant human procathepsin L. Contribution of both intermolecular and unimolecular events in the processing of procathepsin L *in vitro*. *J. Biol. Chem.* **273**, 4478–4484.
7. Mason, R.W. & Massey, S.D. (1992) Surface activation of procathepsin L. *Biochem., Biophys. Res. Commun.* **189**, 1659–1666.
8. Gal, S. & Gottesman, M.M. (1986) The major excreted protein of transformed fibroblasts is an activable acid-protease. *J. Biol. Chem.* **261**, 60–1765.
9. Brix, K., Lemansky, P. & Herzog, V. (1996) Evidence for extracellularly acting cathepsins mediating thyroid hormone liberation in thyroid epithelial cells. *Endocrinology* **137**, 1963–1974.
10. Kakegawa, H., Nikawa, T., Tagami, K., Kamioka, H., Sumitani, K., Kawata, T., Drobnic-Kosorok, M., Lenarcic, B., Turk, V. & Katunuma, N. (1993) Participation of cathepsin L on bone resorption. *FEBS Lett.* **321**, 247–250.
11. Goretzki, L., Schmitt, M., Mann, K., Calvete, J., Chucholowski, N., Kramer, M., Günzler, W.A., Jänike, F. & Graeff, H. (1992) Effective activation of the proenzyme form of the urokinase-type plasminogen activator (pro-uPA) by the cysteine protease cathepsin L. *FEBS. Lett.* **297**, 112–118.
12. Trabandt, A., Aicher, W.K., Gay, R.E., Sukhatme, V.P., Nilson-Hamilton, M., Hamilton, R.T., McGhee, J.R., Fassbender, H.G. & Gay, S. (1990) Expression of the collagenolytic and Ras-induced cysteine proteinase cathepsin L and proliferation-associated oncogenes in synovial cells of MRL/I mice and patients with rheumatoid arthritis. *Matrix* **10**, 349–361.
13. Yagel, S., Warner, A.H., Nellans, H.N., Lala, P.K., Waghorne, C. & Denhardt, D.T. (1989) Suppression by cathepsin L inhibitors of the invasion of amnion membranes by murine cancer cells. *Cancer Res.* **49**, 3553–3557.
14. Ishidoh, K. & Kominami, E. (1998) Gene regulation and extracellular functions of procathepsin L. *Biol. Chem.* **379**, 131–135.
15. Würl, P., Taubert, H., Meye, A., Dansranjavin, T., Weber, E., Günther, D., Berger, D., Schmidt, H., Dralle, H. & Rath, F.W. (1997) Immunohistochemical and clinical evaluation of cathepsin expression in soft tissue sarcomas. *Virchows Arch.* **430**, 221–225.
16. Weber, E., Günther, D., Laube, F., Wiederanders, B. & Kirschke, H. (1994) Hybridoma cells producing antibodies to cathepsin L have greatly reduced potential for tumour growth. *J. Cancer Res. Clin. Oncol.* **120**, 564–567.
17. Coulombe, R., Grochulski, P., Sivaraman, J., Ménard, R., Mort, J.S. & Cygler, M. (1996) Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO J.* **15**, 5492–5503.
18. Fujishima, A., Imai, Y., Nomura, T., Fujisawa, Y., Yamamoto, Y. & Sugawara, T. (1997) The crystal structure of human cathepsin L complexed with E-64. *FEBS Lett.* **407**, 47–50.
19. Turk, D., Guncar, G., Podobnik, M. & Turk, B. (1998) Revised definition of substrate binding sites of papain-like cysteine proteases. *Biol. Chem.* **379**, 137–147.
20. Kärgel, H.J., Dettmer, R., Etzold, G., Kirschke, H., Bohley, P. & Langner, J. (1980) Action of cathepsin L on the oxidized B-chain of bovine insulin. *FEBS Lett.* **114**, 257–260.
21. Kärgel, H.J., Dettmer, R., Etzold, G., Kirschke, H., Bohley, P. & Langner, J. (1981) Action of rat liver cathepsin L on glucagon. *Acta Biol. Med. Ger.* **40**, 1139–1143.
22. Lalmanach, G., Serveau, C., Brillard-Bourdet, M., Chagas, J.R., Mayer, R., Juliano, L. & Gauthier, F. (1995) Conserved cystatin segments as models for designing substrates and inhibitors of cysteine proteinases. *J. Protein Chem.* **14**, 645–653.
23. Gauthier, F., Moreau, T., Lalmanach, G., Brillard-Bourdet, M., Ferrer-Di Martino, M. & Juliano, L. (1993) A new, sensitive fluorogenic substrate for papain based on the sequence of the cystatin inhibitory site. *Arch. Biochem. Biophys.* **306**, 304–308.
24. Schaschke, N., Assfalg-Machleidt, I., Machleidt, W. & Moroder, L. (1998) Substrate/propeptide-derived endo-epoxysuccinyl peptides as highly potent and selective cathepsin B inhibitors. *FEBS Lett.* **421**, 80–82.
25. Gallop, M.A., Barrett, R.W., Dower, W.J., Fodor, S.P.A. & Gordon, E.M. (1994) Applications of combinatorial technologies to drug discovery. 1. Background and peptide combinatorial libraries. *J. Med. Chem.* **37**, 1233–1251.
26. Früchtel, . J.S. & Jung, G. (1996) Organic chemistry on solid supports. *Angew. Chem. Int. ed. Eng.* **35**, 17–42.
27. Furka, A., Sebestyén, F., Asgedom, M. & Dibó, G. (1991) General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Peptide Protein Res.* **37**, 487–493.
28. Houghten, R.A., Pinilla, C., Appel, J.R., Blondelle, S.E., Dooley, C.T., Eichler, J., Nefzi, A. & Ostresh, J.M. (1999) Mixture-based synthetic combinatorial libraries. *J. Med. Chem.* **42**, 3743–3778.
29. Fleckenstein, B., Kalbacher, H., Muller, C.P., Stoll, D., Halder, T., Jung, G. & Wiesmüller, K.-H. (1996) New ligands binding to the human leukocyte antigen class II molecule DRB1*0101 based on the activity pattern of an undecapeptide library. *Eur. J. Biochem.* **240**, 71–77.
30. Wiesmüller, K.-H., Feiertag, S., Kienle, S., Stoll, D., Herrmann, M. & Jung, G. (1996) In *Combinatorial Peptide and Nonpeptide Libraries* (Jung, G.), pp. 203–246. Verlag Chemie, Weinheim.
31. Jung, G. & Beck-Sickinger, A. (1992) Multiple peptide synthesis methods and their applications. *Angew. Chem. Int. ed. Eng.* **31**, 367–383.
32. Stevanovic, S. & Jung, G. (1993) New ligands binding to the human leukocyte antigen class II molecule DRB1*0101 based on the activity pattern of an undecapeptide library. *Anal. Biochem.* **212**, 212–220.

33. Metzger, J.W., Kempter, C., Wiesmüller, K.-H. & Jung, G. (1994) Electrospray mass spectrometry and tandem mass spectrometry of synthetic multicomponent peptide mixtures: determination of composition and purity. *Anal. Biochem.* **219**, 261–277.
34. Müller, C. & Sewald, N. (1998) Synthesis of Fmoc-β-homoamino acids by ultrasound-promoted Wolff-rearrangement. *Synthesis* 837–841.
35. Weber, E., Bahn, H. & Günther, D. (1997) Monoclonal antibodies against cathepsin L and procathepsin L of different species. *Hybridoma* **16**, 159–166.
36. Barrett, A.J. & Kirschke, H. (1981) Cathepsin B, cathepsin H, cathepsin L. *Methods Enzymol.* **80**, 535–561.
37. Chagas, J.R., Ferrer-Di Martino, M., Gauthier, F. & Lalmanach, G. (1996) Inhibition of cathepsin B by its propeptide: use of overlapping peptides to identify a critical segment. *FEBS Lett.* **392**, 233–236.